

# New Approach to Teaching Japanese Pronunciation in the Digital Era Challenges and Practices

Ueyama Motoko

Alma Mater Studiorum, Università di Bologna, Italia

**Abstract** Pronunciation has been a black hole in the L2 Japanese classroom on account of a lack of class time, teacher's confidence, and consciousness of the need to teach pronunciation, among other reasons. The absence of pronunciation instruction is reported to result in fossilized pronunciation errors, communication problems, and learner frustration. With an intention of making a contribution to improve such circumstances, this paper aims at three goals. First, it discusses the importance, necessity, and effectiveness of teaching prosodic aspects of Japanese pronunciation from an early stage in acquisition. Second, it shows that Japanese prosody is challenging because of its typological rareness, regardless of the L1 backgrounds of learners. Third and finally, it introduces a new approach to teaching L2 pronunciation with the goal of developing L2 comprehensibility by focusing on essential prosodic features, which is followed by discussions on key issues concerning how to implement the new approach both inside and outside the classroom in the digital era.

**Keywords** Teaching Japanese pronunciation. Teaching Japanese prosody.

**Summary** 1 Introduction. – 2 Prosody First. – 3 Typologically Unique Characteristics of Japanese Prosody. – 3.1 Rhythmic Component. – 3.2 Tonal Component. – 4 New Approach to Teaching Japanese Pronunciation.



Edizioni  
Ca' Foscari

Ca' Foscari Japanese Studies 13 | Linguistics and Language Education 1

ISSN 2724-1203

ISBN [ebook] 978-88-6969-428-8 | ISBN [print] 978-88-6969-429-5

Peer review | Open access

Submitted 2020-02-14 | Accepted 2020-05-14 | Published 2020-07-03

© 2020 | Creative Commons Attribution 4.0 International Public License

DOI 10.30687/978-88-6969-428-8/010

201

## 1 Introduction

It is not hard to imagine that any dedicated language learner strongly wishes and hopes to be able to perform oral communication successfully in their target language. The acquisition of oral communication is one of the major goals of the communicative approach, the current leading approach to foreign/second language teaching. There is also an increasing demand in linguistic and cultural mediation in speech communication among speakers of various languages, including Japanese. Japanese is one of the most widely spoken languages in the world. It is spoken by more than 128 million people and is ranked thirteenth on the list of world languages by total number of speakers (as L1: 128,229,330; as L2: 121,500).<sup>1</sup> The number of learners of Japanese has been increasing as globalization continues. Japan Foundation, which publishes a survey report on Japanese-language education abroad every three years, released the interim report for its latest survey in 2018. According to the report, 3,850,000 learners of Japanese are spread across 142 countries, and it showed not only an increase in learners (+5.2%) but also in teachers (+20.3%) and institutions (+15%).<sup>2</sup> The number of visitors to Japan is also skyrocketing, having doubled since the beginning of the present decade (124,492 in 2019), according to Japan National Tourism Organization (2019). The number of foreign workers is expected to grow in Japan, which is facing a continuing decline in the labor population. Thus, the acquisition of oral communication is more important than ever.

To learn oral communication, one of the logical initial steps to take is to learn the pronunciation of the target language. Pronunciation training improves speaking abilities by helping learners to develop clear speaking skills that improve intelligibility and minimize effort for interlocutors (Darcy 2018). It also improves learners' perception abilities (Linebaugh, Roche 2015). Successful L2 communication cannot take place without correct pronunciation (Celce-Murcia et al. 2010).

Generally, however, there tends to be very little emphasis on teaching pronunciation in a language classroom (Macdonald 2002), which is also the case for teaching Japanese as a foreign or second language (JFL and JSL, respectively henceforth). There is more than one reason for the lack of pronunciation teaching (Isomura 2001; Toda 2009; Abe et al. 2017). First, there is a belief among teachers that there is no problem in that even though learners have an accent, what they say is understood. Second, it is considered difficult to allocate limited class time to teaching pronunciation. Third, there has been no established method of teaching Japanese pronunciation, and the ma-

---

<sup>1</sup> Ethnologue 2019.

<sup>2</sup> Japan Foundation 2019.

jority of textbooks do not deal with pronunciation systematically. No lexical accent information is provided for new vocabulary except in China (Abe et al. 2017). Fourth, teachers may not feel confident either about their knowledge of Japanese pronunciation or their ability to evaluate students' pronunciation and correct their problems. The latter problem is considered true particularly for non-native-speaking teachers who constitute 70% of the total number of Japanese teachers in the world (Japan Foundation 2015).

The lack of pronunciation instruction could result in learners' difficulties with pronunciation regardless of their L1 backgrounds and proficiency levels, as shown by Toda (2008a; 2009). She conducted a wide-scale survey among 1,216 international students from 47 different countries who enrolled in Japanese pronunciation courses offered by Waseda University (Tokyo, Japan). In their responses, they were able to describe their specific pronunciation errors which they knew how to correct. Their voices also showed how pronunciation problems could hinder not only smooth communication but also learning other linguistic aspects (e.g. vocabulary, listening comprehension). Limited pronunciation skills could lower learners' self-confidence and result in negative effects for learners in estimating their own social credibility and abilities (Morley 1998). This is indeed the case for the participants in the survey as well. They received comments from native speakers, such as that they sounded funny or strange, what they said was not understandable and so on. Such experiences caused frustration, ruining their sense of accomplishment.

It seems reasonable to generalize the problems reported in Toda's study as common among JFL learners in Europe. With the intention of making a contribution to improve the situation, this paper has three goals. The first is to show the importance, necessity, and effectiveness of teaching prosodic aspects of Japanese pronunciation from an early stage of acquisition, which is particularly the case for JFL learners whose L1 backgrounds are major European languages. The second is to provide a panoramic overview of major characteristics of Japanese prosody, which is challenging for learners to acquire because of its typological uniqueness. The third is to introduce a rising approach to teaching Japanese pronunciation in the digital era that is aimed at developing learners' autonomous learning skills.

## 2 Prosody First

What is prosody? The word 'prosody' derives from ancient Greek, where it was used for a "song sung with instrumental music" (Nootboom 1997). Indeed, prosody is sometimes called the "musical" aspect of speech, since it involves variables used to describe music such as pitch contours (melody), rhythm, phrasing, emphasis, timbre (voice

quality), silence (pause), and so on (Truax 2000, 39). In modern phonetics the word ‘prosody’ and its adjectival form ‘prosodic’ are used to refer to those properties of speech extending over larger units of speech than individual sounds or segments. Prosody is also commonly called suprasegmentals.

Frontier approaches to teaching Japanese pronunciation put emphasis on prosody, not on segments, as shown by the dominance of prosody exercises in major textbooks of Japanese pronunciation, as shown in Table 1 that lists the titles of units of *Hitori demo manaberu nihongo no hatsuon* (‘Learn Japanese Pronunciation by Yourself’, Kinoshita, Nakagawa 2019). Units to exercise individual sounds are in gray, occupying only one unit in each textbook. Much more emphasis is placed on prosody than segments for several reasons. The first reason is because Japanese is much more difficult to learn at the prosodic level than at the segmental level. Japanese is one of languages whose phonetic inventory is relatively small. This means that Japanese phonemes are likely to be available in learners’ L1s, which is very likely to result in less probability for errors.<sup>3</sup> This is also the case for five representative European languages – i.e. English, German, French, Italian and Spanish – whose phonemic inventory is larger than that of Japanese, as shown in Table 2.

The second reason is the great importance of prosody in speech communication. Prosody conveys not only a broad range of linguistic information such as lexical information, syntactic information, chunking the stream of speech in phrases, signaling new and contrastive information and disambiguating sentences, but also rich paralinguistic information, i.e. information related to the identity, age, gender, and emotional state of the speaker (Lengeris 2012). Prosodic information present in fluent speech helps the listener perceive the utterances. Words, like musical notes, are grouped together into phrases by their rhythmic and durational properties as well as their tonal pitch. This organization of prosodic phrasing (grouping of words within an utterance) affects the understanding of sentences (Frazier, Carlson, Clifton 2006). The comprehension of spoken language is a complex skill to map the acoustic signals of the speaker’s output onto linguistic units, such as phonemes, syllables, and words

---

**3** As for learners whose L1s are Romance languages, the exception is /h/. In Romance languages, the alphabetic letter *h* is not pronounced (e.g. *homme* /om/ ‘man’ in French; *ha* /a/ ‘have (third person singular)’ in Italian; *hola* /ola/ ‘hello’ in Spanish). This characteristic tends to result in errors such that /h/ is not pronounced when speakers of those languages learn a language with phonemic /h/: e.g. English *hair* and Japanese *hai* ‘yes’ pronounced as *air* and *ai*, respectively. However, note that the availability of phonemes of a target language does not guarantee segmental acquisition, although it tends to facilitate, as shown by /h/ lenition observed among Indonesian learners of Japanese although /h/ is phonemic in Indonesian (Hatasa, Takahashi, Ito 2016).

(Ayasse, Alexis, Wingfield 2018). At the sentence level, prosodic cues help the listener to detect the lexical meanings of the expressions uttered, determine the syntactic structure of the utterance, and comprehend utterance meanings. The importance of prosodic functions is also held in L2 speech communication. In the perception of L2 speech by L1 speakers, prosody plays a more significant role than individual sounds, as reported in earlier research. Studies comparing the relative contribution of segmental vs. prosodic features in degree of foreign accent have shown that deviations in prosodic features may affect listeners' judgement more than deviations in segmental features (Lengeris 2012). Specifically, prosody has been found to be linked to intelligibility, comprehensibility, and accentedness of L2 speech (for a review, among others, see Lengeris 2012). This is also the case for L2 Japanese. A stronger effect has been found not only on native Japanese speakers' evaluation of L2 Japanese pronunciation (Sato 1995) but also on the comprehensibility and naturalness of L2 Japanese pronunciation as perceived by L1 Japanese speakers (Kato et al. 2012; Saito, Akiyama 2017).

**Table 1** Titles of units of *Hitori demo manaberu nihongo no hatsuon* ('Learn Japanese Pronunciation by Yourself', Kinoshita, Nakagawa 2019)<sup>4</sup>

- 1) Introduction – prepare self-learning - how to use OJAD and Praat
- 2) Slash reading 1 – comprehensible and intelligible intonation
- 3) Slash reading 2 – how to express emotions
- 4) Noun and adjective accent – pitch control
- 5) Verb accent – mountain-shaped, plateau-shaped accent and intonation
- 6) Sentence-final intonation 1 – *ka, ne, yo*
- 7) Sentence-final intonation – *janai, yone, kana, kane*
- 8) Rhythm 1 – long vowel – geminate – coda nasal N
- 9) Rhythm 2 – rhythmic patterns of words and – *senryū*
- 10) Vowels and consonants – how to pronounce
- 11) Sound changes – gender and dialect variations
- 12) How to express feelings 1 – politeness
- 13) How to express feelings 2 – roles and characters
- 14) Conclusion – toward future pronunciation learning

**Table 2** The numbers of phonemic consonants and vowels in Japanese, English, German, French, Italian, and Spanish

Languages	Consonants	Vowels
Japanese	17	5
English	24	20
German	23	19
French	20	13
Italian	23	7
Spanish	18	5

The third reason is its overall complex nature, as shown earlier. Prosody consists of multiple components; two essentials are tonal and rhythmic components. In modern linguistic theories, both components are considered hierarchical with different levels of organization (for reviews, see Jun 2005 for intonation; Arvaniti 2009 for rhythm). Every language has prosodic grouping and prosodic prominence at multiple levels (word, phrase, utterance), but different languages use them in very different ways (Jun 2005). Languages differ in their inventory of prosodic units, way of grouping prosodic units (i.e. prosodic grouping), and way of expressing prosodic prominence. To learn and teach Japanese pronunciation, it must be worthwhile and helpful to be familiar with essential characteristics of Japanese prosody.

Lastly, more attention needs to be paid to prosodic aspects since Japanese prosody is particularly challenging to learn regardless of learners' L1 backgrounds, as is well known from both research and practice. Earlier studies have shown that common errors that occur regardless of learners' L1 backgrounds are related to prosody, as opposed to segmental errors that tend to be found among learners whose L1s are specific<sup>5</sup> (Kondo 2011).

### 3 Typologically Unique Characteristics of Japanese Prosody

Why is it so challenging to learn Japanese prosody? It is mainly due to its typologically unique characteristics (Hayashi 2018). Among a wide range of its functions, two key essential elements of Japanese

---

<sup>5</sup> For example, the aforementioned problem of silent /h/ tends to occur only in the production of L2 Japanese by learners whose L1s do not have /h/ (e.g. French, Italian, and Spanish), while this error tends not to occur among learners whose L1s have /h/. Similarly, Korean learners of Japanese have difficulty in pronouncing ザ、ズ、ゼ、ゾ /za, zu, ze, zo/ since /z/ is not a part of the Korean phonemic inventory.

prosody at the lexical level are *lexical pitch accent* and *mora*. Both have lexical functions: i.e. they are used to distinguish meanings of words. Japanese intonation is generated on the basis of the distribution of lexical pitch accent, while rhythm is organized with the use of *mora* as a basic unit. The *mora* is subsyllabic, the smallest of prosodic units of languages. For this, Japanese is classified as a *pitch-accent language* from a pitch or tonal point of view and as a *mora-timed language* from a rhythmic point of view. Japanese is one of the very rare languages classified as a pitch-accent, mora-timed language. This typological uniqueness tends to result in difficulties in learning Japanese prosody for learners whose L1s have very different prosodic characteristics.

The specific elements of L2 Japanese speech that are difficult for a large number of learners to acquire are the properties of Japanese speech that are typologically unique: the phonemic length contrasts for both consonants and vowels and pitch accent. Both are listed as the most common problems by Toda (2009) and Kondo (2011), based on results of their wide-scale surveys.<sup>6</sup> To further understand the nature of learning difficulties, Japanese phonemic length contrasts and pitch accent will be explained mainly from a typological point of view in the rest of this section, considering essential phonological and phonetic characteristics within the rhythmic and tonal components of Japanese speech. Issues related to the L2 acquisition of those properties will be also mentioned.

### 3.1 Rhythmic Component

#### 3.1.1 Mora-timed Temporal Organization

In terms of rhythm or temporal organization, languages are typologically classified into three types: stressed-timed, syllable-timed, and mora-timed (for reviews, among many others, Dauer 1983; Arvaniti 2009). This classification is originally proposed, on the basis of the isochrony hypothesis that claims two points: 1) every language belongs to one particular rhythm type; and 2) rhythm types are defined in terms of a timing unit (syllable, foot, *mora*) that is of equal duration. A number of studies were conducted to test the theory, but ex-

---

<sup>6</sup> Kondo (2011) conducted a survey among 103 Japanese teachers to discover pronunciation problems. Their responses reported pronunciation errors observed among learners from 21 language groups, also showing that common pronunciation errors were those of moraic length contrasts and also moraic coda nasal /N/. These problems were the most common ones that emerged from Toda's (2009) survey study introduced in § 1 of the present paper.

perimental results were too inconsistent to support the theory (Dauer 1983; Arvaniti 2009).

Later studies have proposed that different timing types are characterized rather by the combination of different properties of speech, such as syllable structure, the distribution of word stress/accent, the phonetic realization of lexically prominent syllables (e.g. stressed and pitch-accented syllables in English and Japanese, respectively), the reduction of unstressed/unaccented syllables, and so on (Arvaniti 2009). In stress-timed languages (e.g. English and German), more complex syllable structures are found in stressed syllables, and syllables with more complex structures tend to be stressed (Dauer 1983), and are longer in duration and greater in intensity. Syllable structure and stress are more likely to reinforce each other in a stress-timed than a syllable-timed language. The inventory of syllable types is more limited in syllable-timed languages like French, Italian, and Spanish with less vowel reduction, and even more limited in mora-timed languages like Japanese with neither accentual lengthening nor vowel reduction (e.g. Beckman 1986). The proportion of CV, a syllable consisting of one consonant preceding a vowel, is remarkably higher in Japanese than in English and Spanish, thus showing a smaller proportion due to a wider distribution among different types of syllables. These differences in multiple properties of speech lead to the greatest durational contrasts between stressed and unstressed syllables in stress-timed languages, the smallest durational contrasts between accented and unaccented syllables in Japanese (a mora-timed language that does not have a stress accent system), and somewhere in the middle in syllable-timed languages.

The majority of languages are stress-timed or syllable-timed, and only a few modern languages are classified as mora-timed. Japanese timing patterns are phonetically characterized by the small amount of durational malleability at the prosodic level (Ueyama 2012) mainly due to the absence of lengthening accented syllables and the absence of the vowel reduction of unstressed syllables.<sup>7</sup> These characteristics are difficult to acquire, especially for learners whose L1s are stress-timed languages with a greater amount of durational malleability that is manifested with stressed-syllable lengthening and unstressed-syllable reduction.

---

<sup>7</sup> This does not mean that there is no vowel reduction in Japanese. High vowels /i/ e /u/ tend to be reduced in two phonological contexts: 1) both /i/ e /u/ between voiceless obstruents (e.g. *sushi* [súʃi] 'sushi' where /u/ is not voiced fully); and 2) high back vowel /u/ in sentence-final position (e.g. *ikimasu* [ikimasú] 'go (polite present affirmative form)').



### 3.1.2 Moraic Length Contrasts: Characteristics and L2 Acquisition Issues

Apart from the aforementioned factors, Japanese rhythm is determined largely by phonemic length contrasts. Japanese has two phonemic lengths for consonants as well as for vowels: e.g. *kite* ‘wear (request form)’ vs. *kit:e* ‘postal stamp’ for single vs. geminate consonants; *kite* ‘wear (request form)’ vs. *ki:te* ‘listen (request from)’ for short vs. long vowels. Some languages have only phonemic vowel lengths (e.g. Fijian, Thai, Scottish Gaelic, Swedish, Vietnamese) while some have only consonant lengths (e.g. Italian, Turkish). Only a few languages have both consonant and vowel length contrasts such as Arabic, Estonian, Finnish, Hungarian, and Japanese. This may partly explain why they are very difficult to learn for the great majority of learners.

There has been continuous research on the L2 acquisition of phonemic length contrasts that is still ongoing. A set of general findings emerges from the results of earlier studies (see Hirata 2015 for a comprehensive review). At an early stage of acquisition, L1 backgrounds of learners affect perception significantly. Learners that are not familiar with phonemic length contrasts in their L1s do not perceive Japanese length contrasts with clear and stable categorical boundaries like native Japanese speakers do. Their perception varies depending on phonetic contexts such as lexical pitch accent, speech rate, and positions in a word (e.g. word-initial contrasts are the easiest to perceive), while L1 Japanese perception is stable. In contrast, learners whose L1s have phonemic lengths (e.g. Finnish, Arabic) can perceive Japanese length contrasts in a categorical way similar to that of native Japanese speakers. Vowel length tends to be easier to perceive than consonant length, but it is possible to learn to perceive both types of length contrasts eventually. Production is more challenging than perception: learners improve their production as they advance their study of Japanese, but there always seems to be individual variation in the degree of improvement.

## 3.2 Tonal Component

### 3.2.1 Pitch-accent Language

The world’s languages have been classified into three categories according to their use of pitch at the lexical level (Jun 2005): tone languages, stress (accent) languages, and pitch-accent languages. Japanese belongs to the third category. The exact percentage for each category is not available. However, the World Atlas of Language

Structures/WALS (Haspelmath et al. 2005; Dryer, Haspelmath 2013), i.e. a large database of structural (phonological, grammatical, lexical) properties of languages that were gathered from descriptive materials (such as reference grammars) provides information on word prosody for 176 out of the 200 sample languages included in the database. Since the database is quite representative, we can have an idea of the distribution of the three types: 141 (80%) use stress, and 28 (16%) have only lexical tone or pitch accent (Goedemans 2010). This information indicates that pitch-accent languages such as Japanese are the minority.

The majority of tone languages are spoken in Asia, such as Mandarin Chinese, Thai, and Vietnamese and also in Africa, such as Akan, Igbo, and Yoruba. Each tone language has its own inventory of lexical tones (contour tones and level/register tones in Asia and Africa, respectively) applied to a syllable that are phonologically distinctive by showing unique pitch patterns to convey different meanings. In contrast, in both stress (accent) languages and pitch-accent languages, only one syllable or mora in a word is more prominent than the others. A major difference between lexical stress and pitch accent<sup>8</sup> is in which acoustic correlate it is involved. In stress languages, word stress involves multiple acoustic parameters as in English where word stress is produced with changes in fundamental frequency (F0), intensity, duration, vowel quality, and so on (among others, Beckman 1986). In pitch-accent languages, lexical pitch accent is achieved principally by pitch change. Other than Japanese, major modern languages that have been identified as pitch-accent languages are Serbo-Croatian, Slovenian, Norwegian, Persian, Punjabi, Swedish, Western Basque, and certain dialects of Korean (among others, Jun 2005; Hyman 2006). Some have both pitch accent and stress, such as Serbo-Croatian, Norwegian, and Swedish, but Japanese has only pitch accent. Some have both lexically accented and accentless words, such as Japanese and Northern Bizkaian Basque. Lastly, some languages do not have any lexical tones, stress or accent; French and Seoul Korean belong to this category. In these languages, intonation patterns are determined only based on post-lexical tones (Jun 2005).

As mentioned earlier, a major difference between pitch accent and stress (accent) is that only pitch is involved for the former while multiple correlates exist for the latter. This difference explains one of

---

<sup>8</sup> In a current common intonation model, the autosegmental-metrical (AM) approach (among others, see Ladd 1996; Jun 2005), the term *pitch accent* is used also for post-lexical prominence that is assigned to a word in a certain speech context. Thus, there are two types of pitch accent: lexical accent and post-lexical pitch accent. Post-lexical pitch accents do not have distinctive functions without changing the lexical identity of the word, opposed to lexical pitch accents that are used to distinguish meanings, as in Japanese.

the common problems among learners whose L1s are stress (accent) languages; they tend to use their L1 stress in place of the Japanese of pitch accent by lengthening vowels and/or using more intensity at an early stage of acquisition, as shown in past experimental studies (e.g. Ueyama 2012 for L2 Japanese-L1 English; Asano, Gubian 2018 for L2 Japanese-L1 German; Ueyama 2016 for L2 Italian-L1 Japanese). Since Japanese has phonemic vowel lengths, non-native extra durational stretch native Japanese speakers' perception of (C)V as (C)VV: e.g. /waka'ru/ 'understand' is perceived as /wakaaa'ru/ with long vowel /aa/. Learners whose L1s do not have a lexical accent face a different problem that is also caused by a difference between their L1s and Japanese. They simply use their L1 tonal shapes at the sentence level.

### 3.2.2 Japanese Lexical Accent: Characteristics and L2 Acquisition Issues

The lexical or word accent system of Tokyo Japanese<sup>9</sup> is characterized by the following principal properties. There is only one type of pitch accent, HL,<sup>10</sup> a high tone followed by a falling tone. The presence, absence, and position of pitch accent HL are contrastive (e.g. McCawley 1968, and many others). Only one HL pitch fall is allowed within a word, and pitch cannot rise again within the same word once it goes down: i.e. there can maximally be one prominence within a word (Kawahara 2015).

The following steps or rules are applied to have final tonal shapes. If there is no lexical accent on the first mora of the word, pitch rises from low to high from the first onto the second mora of a phrase-initial word (i.e. phrase-initial pitch rise), and pitch stays high up to the lexical accent. The three tonal patterns of two-mora words with three distinctive meanings are presented in [fig. 1]: *ha\*shi-da* 'they are chopsticks'; *hashi\*-da* 'it's a bridge'; *hashi-da* 'it's an edge' (the underlined syllable is lexically accented while the asterisk marks the approximate location of pitch fall). These three tonal patterns can be abstractly represented by a sequence of high and low tones: H\*LL,

---

<sup>9</sup> Different Japanese dialects are characterized by varying characteristics of different properties of speech, including word accent.

<sup>10</sup> H\*+L is alternatively used to represent lexical pitch accent mainly in research works conducted with the AM approach as well as in the J-ToBI prosodic labeling scheme (Venditti 1997; for the extended version of J-ToBI, Maekawa et al. 2002).

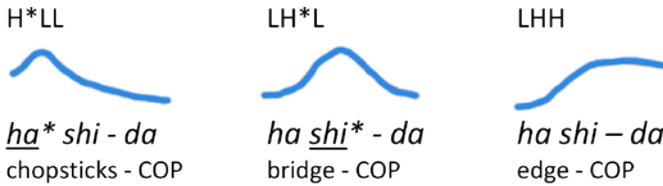


Figure 1 An example of a minimal set of two-mora words distinguished by pitch accent

---

LH\*L, LHH.<sup>11</sup> In this way, a tonal pattern of one-word unit, *bunsetsu*,<sup>12</sup> can be explained or predicted by a small set of steps once the presence, absence, and position of pitch accent HL are found.

### 3.2.3 The L2 Acquisition of Japanese Lexical Pitch Accent

Lexical pitch accent is another element, along with phonemic length contrasts, that causes difficulty for learners of Japanese; L2 acquisition of lexical pitch accent has been studied extensively up to today (for reviews, see Hirata 2015; Hatasa, Takahashi, Ito 2016). Main findings that emerged from past studies are as follows. The degree of perceptual accuracy for different pitch accent patterns and effects of syllable structures depends on learners' L1s. The learners show higher accuracy in those pitch accent patterns that are similar to the prosodic patterns of their L1s, but accentless words and words with word accent on the last mora (e.g. LHH and LHH\*, respectively, for three-mora words with no word-internal pitch fall) are perceived more correctly than words with a pitch fall within a word (e.g. H\*LL and LH\*L).

Perceptual accuracy varies across individuals, regardless of their history of Japanese language study. Even advanced speakers cannot reach close to a native level of perceptual accuracy, unlike the case of phonemic length contrasts that advanced speakers are able to perceive with native-like accuracy. A similar tendency is found also in

---

**11** Using a binary system of H and L tones has been criticized since it is considered not to represent phonological and phonetic characteristics of Japanese accent patterns, and these days, it is more common to indicate only the position of pitch accent. However, the binary system is employed in this paper for the ease of explanations of L2 patterns in this section.

**12** *Bunsetsu* is a morphological unit of accentuation consisting of a content word such as nouns, verbs, or adjectives with or without being followed by a string of function morphemes such as particles or postpositions, as defined by Hattori (cited in Nagano-Madsen 2015, 200).

production (among others, Sakamoto 2010 and Muradás-Taylor 2019 for English-speaking learners; Pappalardo 2017 for Italian-speaking learners). Commonly, the same learners produce more than one accent pattern for the same lexical item (Ueyama 2012; Hatasa, Takahashi, Ito 2016), which strongly reflects the nature of interlanguage that is characterized by non-systematic variability (Selinker 1972).

Similar to the acquisition of phonemic length contrasts, production seems to lag behind perception (Hirata 2015; Hatasa, Takahashi, Ito 2016). This general tendency is shown by Sakamoto (2010) who investigated both the production and perception of the same learners. “The experienced learners with an average of 3.7 years of Japanese study with 1 year of stay in Japan showed perception of the three types of pitch accent similar to that of NJs, but their production was still significantly lower than that of native speakers” (cited in Hirata 2015, 737). Difficult patterns to learn do not seem to be the same in production and perception, as pointed by Hatasa and colleagues (2016), based on the results of their comparison of findings of earlier research on the acquisition of Japanese pitch accent by English-speaking learners.

Lexical pitch accents are fundamental components of the phrase-level intonation of Japanese. Japanese intonation can be largely determined at the lexical level based on the distribution of lexical pitch accents and interacting with post-lexical processes such as phrase-initial pitch rise (see also § 3.2.2) and downstep<sup>13</sup> (Pierrehumbert, Beckman 1988; Venditti 1997; Maekawa et al. 2002; Jun 2005; Igarashi 2015). The lack of acquiring lexical pitch accents at the production level could result not only in misunderstandings of meanings of words but also in unnatural intonation at the sentence and discourse level, which may interfere with smooth communication by reducing the comprehensibility of learners’ speech.

#### 4 New Approach to Teaching Japanese Pronunciation

The overview of the essential characteristics of Japanese prosody mainly at the lexical level has shown two points: 1) common difficulties in learning prosodic features are largely due to their typological uniqueness; and 2) such difficulties cannot be easily overcome by adding experiences of learning Japanese with no pronunciation instructions. Under these circumstances, the logical solution is to

---

**13** Downstep is a phonological process in which the local pitch height of each accental phrase typically consisting of one lexical word plus any following particles (Igarashi 2015), i.e. *bunsetsu*, is reduced when followed by a lexically accented phrase, which results in forming a staircase-like effect of accental phrase heights in sequence (see Jun 2005).

teach Japanese pronunciation from an early stage of learning. Then the next question is how? The following points are keys to answering the question.

The first point is a general goal: i.e. teach *comprehensible* pronunciation. In the last decade, there have been on-going shifts in pedagogical implications for pronunciation teaching (Levis 2005, cited in Saito 2018). The traditional approach focuses equally on all L2 pronunciation features to teach native-like accurate pronunciation. The new approach instead focuses selectively on certain features affecting comprehensibility or intelligibility with a goal of teaching comprehensible L2 pronunciation based on the fact that many successful L2 speakers remain accented but highly comprehensible (Saito 2018). The goal of the new approach is also the basis of ongoing efforts to improve the problematic situation of teaching Japanese pronunciation in a classroom.

Second is the need to teach essential prosodic features in a classroom. This line of instruction was proven to be effective in Oyama's (2014) experimental classroom study. Eight sessions of 20 minutes were carried out for one month, focusing on selected prosodic features such as rhythm and mora length contrast, major characteristics of lexical accent patterns (e.g. pitch accent, accent type, compound accent, simple word accent) and those of intonation. In every session, students also practiced dialogues by focusing on accentual phrases. The comparison of results of pre- and post-tests showed the significant effects of instructions. As pointed out in the conclusion of the study, the important function of this type of training is to provide learners with metalinguistic knowledge about Japanese phonology and phonetics. As reported by Toda (2008b, 2008c), successful learners have learned such knowledge and utilize it to monitor their pronunciation critically.

The third key issue is that teachers need to continue to follow up on instructions in classrooms, especially for pitch accent. Teachers are expected not only to continue to teach accent patterns of new vocabularies but also to have the ability to assess learners' pronunciation, explain problems, and carry out exercises to solve problems. However, it is evident that not all teachers have such an ability, including native Japanese teachers. Byun (2018) conducted a test of the perception of accent patterns among 126 Japanese students studying Japanese language education, and results showed that the percentage of correct answers ranged from 60 to 80%. The same study also showed the effects of a training that improved the percentage up to 90%. Similar difficulties are faced also by native Japanese teachers, as shown in Kanamura (2020)'s survey conducted among 69 teachers working in Japan that has unveiled the psychological block to teaching pronunciation among teachers. Kanamura expresses the importance and need to support Japanese teachers to improve their ability

to assess the pitch of one's own voice and understand the difference between model and actual pronunciation.

The fourth point is to utilize digital resources, including teaching materials such as audio files, videos, websites ranging from pages designed to teach Japanese pronunciation (e.g. つたえるはつおん *Tsutaeru hatsuon*)<sup>14</sup> to digital tools to support pronunciation learning online or offline. The most frequently used tool worldwide may be the web-based system OJAD<sup>15</sup> (Online Japanese Accent Dictionary) developed by Minematsu and colleagues. A prosodic reading tutor Suzuki-kun, one of four OJAD features, is utilized not only by learners but also by teachers. Suzuki-kun visualizes intonation curves with pitch accent information for any given text, also generating speech models of a selected voice at three different speech rates with the use of speech synthesis technologies (Minematsu, Hirano, Nakamura 2018). The combination of a visual display of prosodic information and an audio model has proved to be very effective in improving the naturalness of L2 Japanese pronunciation, as shown by experimental evidence (Minematsu et al. 2016). Speech analysis tools (e.g. Praat,<sup>16</sup> WASP,<sup>17</sup> although not developed originally for teaching second language pronunciation, can be utilized by learners to check their pronunciation by comparing with audio models for acoustic patterns. Kinoshita and Nakagawa (2019) propose combining OJAD and Praat. This is currently one of the cutting-edge methods that cover three steps of learners' autonomous learning: i.e. visualize prosodic patterns, generate audio models with OJAD, and check learners' pronunciation by comparing with audio models for acoustic patterns with Praat.

To conclude, this paper has argued for the importance, necessity, and effectiveness of teaching Japanese pronunciation from an early stage of leaning, especially prosodic features, and then it has shown from a typological point of view why it is very difficult to learn Japanese prosody for the majority of learners, regardless of their L1 backgrounds. Last but not least, the new approach to teaching L2 pronunciation with a goal of developing L2 comprehensibility has been introduced, along with key issues concerning how to implement the approach both inside and outside the classroom in the digital era.

---

**14** <http://www.japanese-pronunciation.com/>.

**15** <http://www.gavo.t.u-tokyo.ac.jp/ojad/eng/pages/home>.

**16** <http://www.fon.hum.uva.nl/praat/>.

**17** <https://www.speechandhearing.net/laboratory/wasp/>.

## Bibliography

- OJAD (2016) = Online Japanese Accent Dictionary. URL <http://www.gavo.t.u-tokyo.ac.jp/ojad/eng/pages/home>.
- Abe Shin 阿部新 et al. (2017). “Ōshū ni okeru nihongo kyōiku jijō – kyōshi o taishō to shita ankēto no kekka kara 欧州における日本語音声教育事情 — 教師を対象としたアンケートの結果から— 欧州における日本語音声教育事情 — 教師を対象としたアンケートの結果から (The Current Situation of Teaching Japanese Pronunciation in Europe – The Findings of a Questionnaire Survey by Japanese Teachers)”. *Japanese Education in Europe*, 21, 436-8. URL <https://www.eaje.eu/ja/symposium/31>.
- Arvaniti, Amalia (2009). “Rhythm, Timing and the Timing of Rhythm”. *Phonetica*, 66, 46-63.
- Asano, Yuki; Gubian, Michele (2018). “‘Excuse Meeee!!’: (Mis)coordination of Lexical and Paralinguistic Prosody in L2 Hyperarticulation”. *Speech Communication*, 99, 183-200.
- AYASSE, Nicole; Alexis, Johns; Wingfield, Arthur (2018). “Speech Comprehension and Cognition in Adult Aging”. *Oxford Research Encyclopedia of Psychology and Aging*. DOI <https://doi.org/10.1093/acrefore/9780190236557.013.368>.
- Beckman, Mary (1986). *Stress Accent and Non-Stress Accent*. Berlin: De Gruyter Mouton.
- Boersma, Paul; Weenink, David (2018). *Praat: Doing Phonetics by Computer, Version 6.0.37*. URL <http://www.praat.org/>.
- Byun Hi-Gyung 邊姫京 (2018). “Nihongobogowasha no Tōkyō akusento kiki-tori nōryoku 日本語母語話者の東京語アクセント聞き取り能力 (Tōkyō Accent Perception Ability for Native Speakers from Different Areas)”. *Journal of the Phonetic Society of Japan*, 22, 1-21.
- Celce-Murcia, Marianne et al. (2010). *Teaching Pronunciation: A Course Book and Reference Guide*. New York: Cambridge University Press.
- Darcy, Isabelle (2018). “Powerful and Effective Pronunciation Instruction: How Can We Achieve It?”. *CATESOL Journal*, 30, 13-45.
- Darwin, Chris (1975). “On the Dynamic Use of Prosody in Speech Perception”. Cohen, Antonie; Nooteboom, Siete G. (eds), *Structure and Process in Speech Perception = Proceedings of the Symposium on Dynamic Aspects of Speech Perception held at I.P.O.* (Eindhoven, Netherlands, August 4-6). Berlin: Springer, 178-94.
- Dauer, Rebecca (1983). “Stress-Timing and Syllable-timing Reanalyzed”. *Journal of Phonetics*, 11, 51-62.
- Derwing, Tracey; Munro, Murray (2005). “Second Language Accent and Pronunciation Teaching. A Recent Research Approach”. *TESOL Quarterly*, 39(3), 379-97.
- Dryer, Matthew; Haspelmath, Martin (eds) (2013). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URL <https://wals.info>.
- Ethnologue (2019). *What Are the Top 200 Most Spoken Languages?*. URL <https://www.ethnologue.com/guides/ethnologue200>.
- Frazier, Lyn; Carlson, Katy; Clifton, Charles (2006). “Prosodic Phrasing is Central to Language Comprehension”. *Trends in Cognitive Sciences*, 10(6), 244-9.



- Goedemans, Rob (2010). "A Typology of Stress Patterns". Van der Hulst, Harry; Goedemans, Rob; van Zanten, Ellen (eds), *A Survey of Word Accentual Systems in the Languages of the World*. Berlin: Mouton de Gruyter, 647-66.
- Haspelmath, Martin et al. (eds) (2005). *The World Atlas of Language Structures*. Oxford: Oxford University Press.
- Hatasa Yukiko 畑佐由紀子; Takahashi Eriko 高橋恵利子; Itō Katsuhiro 伊東克洋 (2016). "Dainigengo to shite no nihongo no onseishūtoku to hyōkakenkyū 第二言語としての日本語の音声習得と評価研究 (Acquisition and Assessment of L2 Pronunciation Skills in Japanese Learners)". *Bulletin of the Graduate School of Education, Hiroshima University*. Part. II, *Arts and Science Education*, 65, 177-86.
- Hayashi Ryoko 林良子 (2018). "Nihongo ni okeru inritsushūtoku • kyōiku no kadai to gakushūsha onsei ākaibu no kōchiku 日本語における韻律習得・教育の課題と学習者音声アーカイブの構築 (Problems of Teaching and Learning Japanese Prosody and Development of L2 Japanese Speech Corpus)". *Journal of Kokugakuin University*, 119(11), 82-96.
- Hirata, Yukari (2015). "L2 Phonetics and Phonology". Kubozono, Haruo (ed.), *Handbook of Japanese Phonetics and Phonology*. Berlin: de Gruyter Mouton, 719-62.
- Huckvale, Mark (2016). *WASP (Waveform Spectrogram & Pitch) Version 1.2*. URL <https://www.speechandhearing.net/>.
- Hyman, Larry (2006). "Word-prosodic Phonology". *Phonology*, 23, 225-57.
- Igarashi, Yosuke (2015). "Intonation". Kubozono, Haruo (ed.), *Handbook of Japanese Phonetics and Phonology*. Berlin: de Gruyter Mouton, 525-68.
- Isomura Kazuhiro 磯村一弘 (2001). "Kaigai ni okeru nihongo accent kyōiku no genjō 海外における日本語アクセント教育の現状 (Japanese Accent Education Abroad)". Paper presented in the autumn meeting of the Association of Japanese Language Education.
- Japan Foundation (2015). *Survey Report on Japanese-Language Education Abroad 2015*. URL <https://www.jpf.go.jp/e/project/japanese/survey/result/survey15.html>.
- Japan Foundation (2019). *Survey Report on Japanese-Language Education Abroad 2018* (Breaking News). URL <https://www.jpf.go.jp/j/about/press/2019/029.html>.
- Japan National Tourism Organization (2019). *Trends in Visitor Arrivals in Japan*. URL <https://statistics.jnto.go.jp/en/graph/#graph--inbound-travelers--transition>.
- Jun, Sun-Ah (2005). "Prosodic Typology". Jun, Sun-Ah (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press, 430-58.
- Kanamura Kumi 金村久美 (2020). "Nihongo akusento kyōiku o shien suru tame no kyōiku towa - nihongo kyōin no ankēto chōsa kara 日本語アクセント指導を支援するための音声学教育とは - 日本語教員へのアンケート調査から (Improvements to Phonetic Education That Would Aid in Teaching Japanese Pitch Accent: Suggestions from a Questionnaire Survey of Japanese Language Teachers)". *Journal of the Phonetic Society of Japan*, 24, 36-48.
- Kato, Shuhei et al. (2012). "Effects of Learners' Language Transfer on Native Listeners' Evaluation of the Prosodic Naturalness of Japanese Words". *Proceedings of the 6th International Conference on Speech Prosody, SP 2012*, 198-201.

- Kawahara, Shigeto (2015). "Phonology of the Japanese Accent". Kubozono, Haruo (ed.), *Handbook of Japanese Phonetics and Phonology*. Berlin: de Gruyter Mouton, 445-92.
- Kinoshita Naoko 木下直子; Nakagawa Chieko 中川千恵子 (2019). *Hitoridemo manaberu nihongo no hatsuon-OJAD de shirabete Praat de tashikameyō* ひとりでも学べる日本語の発音—OJADで調べてPraatで確かめよう (Learn Japanese Pronunciation by Yourselfs - Find with OJAD and Check with Praat). Tōkyō: Hitsuji Shobō.
- Kondō Mariko 近藤真理子 (2011). "Nihongo gakushūsha no onsei shūtoku ni okeru dai'ichigengo tokuyū no kanshō to fuhengengoteki kanshō - nihongo kyōshi e no ankēto kekka kara 日本語学習者の音声習得における第一言語特有の干渉と普遍言語的干渉日本語教師へのアンケート調査から (Language Specific and Language Universal L1 Interference on Phonological Acquisition of Japanese: Results of Questionnaires to Japanese Teachers)". *Bulletin of the Graduate Division of Literature of Waseda University*, 3(57), 21-34.
- Ladd, Robert D. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.
- Lengeris, Angelos (2012). "Prosody and Second Language Teaching: Lessons from L2 Speech Perception and Production Research". Romero-Trillo, Jesús (ed.), *Pragmatics and Prosody in English Language Teaching*. New York: Springer, 25-40.
- Linebaugh, Gary; Roche, Thomas (2015). "Evidence That L2 Production Training Can Enhance Perception". *Journal of Academic Language and Learning*, 9, A1-17.
- Macdonald, S. (2002). "Pronunciation - Views and Practices of Reluctant Teachers". *Prospect*, 17(3), 3-15.
- Maekawa et al. (2002). "X-JToBI: An Extended J-ToBI for Spontaneous Speech". *Proceedings of the 7th International Conference on Spoken Language Processing, ICSLP 2002*, 1545-8.
- Minematsu et al. (2016). "Improvement of Naturalness of Learners' Spoken Japanese by Practicing with the Web-based Prosodic Reading Tutor, Suzuki-kun". *Proceedings of the 6th International Conference on Speech Prosody, SP 2016*, 257-61.
- Minematsu, Nobuaki; Hirano, Noriko; Nakamura, Noriko (2018). "Development of OJAD and Its Effectiveness To Improve the Naturalness of Learners' Spoken Japanese". Ueyama, Motok; Srdanović, Irena (eds), *Digital Resources for Learning Japanese*. Bologna: Bononia University Press, 13-34.
- Morley, Joan (1998). "Trippingly on the Tongue: Putting Serious Speech/pronunciation Instruction Back in the TESOL Equation". *ESL Magazine*, 1(15), 20-3.
- Muradás-Taylor, Becky (2019). "'English Japanese': Why Every English Speaker Produces Japanese Pitch Accent Differently and No One Uses Standard Japanese". Paper presented at The 9th International Symposium on the Acquisition of Second Language Speech.
- Nagano-Madsen, Yasuko (2015). "Intonation in Okinawan". Heinrich, Patrick et al. (eds), *Handbook of the Ryukyuan Languages*. Berlin; New York: de Gruyter Mouton, 199-226.
- Nootboom, Sieb (1997). "The Prosody of Speech: Melody and Rhythm". *The Handbook of Phonetic Sciences. Blackwell Handbooks in Linguistics*. Cambridge (MA): Blackwell, 640-73.

- Ōyama Rie 大山理恵 (2014). “Nihongo gakushūsha ni okeru akusentoku no shūtoku – onsei shidō o jūshishita jugyō jissen no kōka 日本語学習者におけるアクセント句の習得 – 音声指導を重視した授業実践の効果 (Acquisition of Japanese Accent Phrase by Japanese Language Learners – Effects of the Japanese Teaching Method by Focusing on Pronunciations)”. *Bulletin of Doshisha University Center for Japanese Language & Cultures*, 14, 91-104.
- Pappalardo, Giuseppe (2017). “Chūkan gengo on'inron kenkyū – itariago bogowasha ni yoru nihongo pitchi akusento no chikaku to seisei no jikken kekka bunseki 中間言語音韻論研究-イタリア語母語話者による日本語ピッチアクセントと知覚と生成の実験結果分析 (Interlanguage Phonology: Experimental Analysis of Perception and Production of Japanese Pitch Accent by Italian Learners of Japanese)”. *The Proceedings of the 20th Japanese Language Symposium in Europe, The Association of Japanese Language Teachers in Europe (AJE)*, 21, 141-7.
- Pierrehumbert, Janet; Beckman, Mary (1988). *Japanese Tone Structure*. Boston: Massachusetts Institute of Technology Press. Linguistic Inquiry Monograph, 15.
- Saito, Kazuya (2012). “Effects of Instruction on L2 Pronunciation Development: A Synthesis of 15 Quasi-Experimental Intervention Studies”. *TESOL Quarterly*, 46, 842-54.
- Saito, Kazuya (2018). *Teaching “Comprehensible” Second language Pronunciation: A Research Synthesis*. Invited talk at the University of Reading.
- Saito, Kazuya; Yuka Akiyama (2017). “Linguistic Correlates of Comprehensibility in Second Language Japanese Speech”. *Journal of Second Language Pronunciation*, 3, 199-217.
- Sakamoto, Emi (2010). *An Investigation of Factors Behind Foreign Accent in the L2 Acquisition of Japanese Lexical Pitch Accent By Adult English Speakers* [PhD dissertation]. University of Edinburgh.
- Sato, Tomonori (1995). “A Comparison of Phonemes and Prosody in the Evaluation of Spoken Japanese”. *Japanese-Language Education Around the Globe*, 5, 139-54.
- Selinker, Larry (1972). “Interlanguage”. *International Review of Applied Linguistics*, 10, 209-31.
- Shport, Irina (2016). “Training English Listeners to Identity Pitch-Accent Patterns in Tōkyō Japanese”. *Studies in Second Language Acquisition*, 38, 739-69.
- Toda Takako 戸田貴子 (2008a). “Nihongo gakushūsha no onsei ni kansuru mondaiten 日本語学習者の音声に関する問題点 (Problems Relating to Pronunciation by Learners of Japanese)”. Toda Takako (ed.), *Nihongo Kyōiku to Onsei 日本語教育と音声 (Japanese Language Education and Pronunciation)*. Tōkyō: Kuroshio Shuppan, 23-42.
- Toda Takako 戸田貴子 (2008b). “Otona ni nattekaramo hatsuon no shūtoku wa kanōka 大人になつてからも発音の習得は可能か (Can Adult Learners Master Pronunciation?)”. Toda Takako (ed.), *Nihongo Kyōiku to Onsei 日本語教育と音声 (Japanese Language Education and Pronunciation)*. Tōkyō: Kuroshio Shuppan, 43-60.
- Toda Takako 戸田貴子 (2008c). “‘Hatsuon no tatsujin’ to wa dono yōna gakushūsha ka 「発音の達人」とはどのような学習者か (What Type of Learners Master Pronunciation?)”. Toda Takako (ed.), *Nihongo Kyōiku to Onsei 日本語教育と音声 (Japanese Language Education and Pronunciation)*. Tōkyō: Kuroshio Shuppan, 61-82.

- Toda Takako 戸田貴子 (2009). “Nihongo kyōiku ni okeru gakushūsha onsei no kenkyū to onsei kyōiku jissen 日本語教育における学習者音声の研究と音声教育実践 (Study of L2 Japanese Pronunciation and Practice of Pronunciation Education)”. *Japanese Language Education*, 142, 47-57. DOI [https://doi.org/10.20721/nihongokyoiku.142.0\\_47](https://doi.org/10.20721/nihongokyoiku.142.0_47).
- Truax, Barry (2000). *Acoustic Communication*. 2nd ed. Westport: Praeger Publishers.
- Ueyama, Motoko (2012). *Prosodic Transfer: An Acoustic Study of L2 English and L2 Japanese*. Bologna: Bononia University Press.
- Ueyama, Motoko (2016). “Prosodic Transfer in L2 Relative Prominence Distribution: The Case Study of Japanese Pitch Accent Produced by Italian Learners”. *Proceedings of the 8th International Conference on Speech Prosody, SP 2016*, 602-60.
- Venditti, Jennifer (1997). “Japanese ToBI Labeling Guidelines”. *OSU Working Papers in Linguistics*, 50, 127-62. URL [https://kb.osu.edu/bitstream/handle/1811/81780/WPL\\_50\\_July\\_1997\\_127.pdf](https://kb.osu.edu/bitstream/handle/1811/81780/WPL_50_July_1997_127.pdf).